Semantic Matching Using Ontology in Multilingual Environment

Makrant I., Mansi M.

Abstract: The tremendous increase in usage of data over the past few years and the ease of availability of things across the globe any time lead to the Advancement of multilingual database. Storage, retrieval and archiving of data for multi lingual system has been a challenge. Our research project "Semantic Matching using Ontology in Multilingual Environment" is an extension is to look into addressing multilingual data. The report focusses on providing design and implementation of multilingual system. These comprise of two main components being (i) Cross Lingual Information Retrieval, and (ii) Indian Language to Indian Language Machine Translation. Consider the context of large scale natural language processing applications in the areas of Cross Lingual IR and Machine Translation, wherein such a model for multilingual dictionary is established. When contrasted to traditional single lingual or bilingual dictionary, the model uses the core concept of Synonym Groupings (synsets) that is used as a way to connect different languages in a crisp and efficient manner.

Keywords: Multilingual database, Storage, Retrieval, Hierarchy, multilingual dictionary, Translation, Algorithm

I. INTRODUCTION

The primary goal of the system is to help the people to get familiar with expressions in everyday conversations and typical dialogues in the target languages (Marathi, Hindi) that can take place in public places like a shopping mall, a railway station, a bus-stop, a restaurant, a post office, a bank etc in a systematic way. To help them better understand the spirit of the city an idea of the most used dialogues in Marathi can be helpful.

Multilingual Database is a database in three languages – Hindi, Marathi and English. It is a search and retrieval system for digital Documents. With the increasing integration of global economy and proliferation of languages other than English into information systems, capability to store and manage data in multiple languages simultaneously is of vital importance. A hierarchical tree structure called 'Ontology' is used. Applications that are catered to domain-specific knowledge being to use Ontology extensively. This is used for semantic matching.

The existing system has following features:

- a. Ontology is in English.
- b. Synsets are available in different languages.

Computational Framework for the Multilingual Dictionary: For effective implementation of our idea of synset based multilingual dictionary, we carefully designed the dictionary development process, which is, in fact,

expected to involve a number of human lexicographers. Figure below shows the complete semi-automatic data flow in the dictionary development process.





II. PROBLEM STATEMENT WITH SCOPE & OBJECTIVE

The customer opens the web-page. He will select the transaction he wants to perform. There are two transactions to be performed. In case of INTER-LANGUAGE TRANSLATION, the user will select a source language and a target language. Then he will enter a word from the database to be translated. In case of SEARCHING OPERATION, he can either input a certain price range and get the products in that range, or can search a certain brand in the hierarchy. *The important factors considered are:* whether methods used are fast and convenient? *and* is the database complete for the shop?

III. PROPOSED SYSTEM

Our proposed system will include: a) Ontology in 2 other languages: Hindi and Marathi. b) Semantic matching in the three languages for words.

An ontology is a shared conceptualization of knowledge in a particular domain. There is need to separate the knowledge about the target domain from the rest of the application code. Ontology helps to facilitate the above. The key benefits of this approach are: simplification of the application code, possible sharing of knowledge among multiple applications, and the flexibility of evolving the knowledge without requiring changes to the application. This paper addresses the problem of supporting ontologybased semantic matching in RDBMS.

Word-Alignment in the Proposed Model: The problem of correct lexical choice on the target language side was tackled by proposing a novel approach of word-alignment across the synsets of languages. Word alignment refers to the mapping of each member of the synset with the most appropriate member of the synset of another language. Consider the example wherein the word under selection is $\nabla(\vec{s}, \vec{\phi})$ 'boy' in Hindi. In this case of 'a young male person' needs to be lexically transferred to Marathi, there are four choices available in the synset, as illustrated in figure2.



Fig.2 Illustration of aligned synset members for the concepts: a youthful male person

Considering Hindi as the pivot, we proposed that each of the four words in Marathi synset be linked to the appropriate Hindi word in the direction Marathi I Hindi and each of the two words in English synset has to be linked with appropriate Hindi word in the direction English I Hindi. As a result, the first and the third member of the Marathi synset(i.e. मुलगा and पोर) are mapped to two different Hindi words(मुलगा -> लड़का, पोर -> बच्च). The second and the fourth member in Marathi synset are linked to one word (पोरगा -> छोकरा and पोरगे -> छोकरा) in the Hindi synset. Three words in Hindi synset(i.e. HW4, HW5, HW7) are left without being linked, as seen in the figure 2. In a situation, when a Marathi word is aligned with a single Hindi word(eg मुलगा -> लड्का for a particular concept in the direction of Marathi to Hindi, from our past experience we assume that the lexical transfer in the reverse direction also holds good, yielding लड्का -> मुलगा.

IV. PROJECT DESIGN

1) System design: The flow diagram is illustrated in the figure3. NetBeans IDE 6.7.1 is used for the Front End. Glassfish server is used to connect the JSP and the database in MS Access.



- 2)Algorithms :
- a)Specific algorithm:
- 1-Choose source language
- 2-Find hierarchy of key word in source language
- 3-Start from the ROOT node.
- 4-Search among the children.
- 5-If the keyword selected by the user is found, display its children in the next dropdown for the next selection.
- 6-If translation is selected as the operation, assign it as the curr_node_level
- 7-Now we search hierarcy level in target language

A- If (curr_node_level = target level1)

Target language Translate

- B- Else If (curr_node_level = target_level2 : leaf node)
- Target language Translate
 - C- Else If (curr_node_level = synset of source_level2 : leaf node)

Target language Translate

D-Word not found in database

b)General algorithm :

Else

i)To find a match between Hindi & English synsets: for each synset identity english_synset_id in English Ontology do if (english_synset_id == hindi_synset_id) do for each relation r pointed by hindi_synset_id do if (relation type of r is semantic) do clamp the synset identity linked by relation r in to english synset id end if else clamp the synset identity linked by relation r in hindi_synset_id to english_synset_id AND manually insert the corresponding lexical element end else end for end if end for

Eg.: Here we have hindi input to English output

The example is as follows :

Hindi word	En	glish Word
दुकान		shop

Take the word **द्वा**.

Consider the id number of the word. Now search for the id number in the English set. If match is found, display the output. Else manually insert the word in the table with an id number.

ii)To find a match between Hindi & Marathi synsets: for each synset hindi identity marathi_synset_id in Marathi Ontology do if (marathi_synset_id == hindi_synset_id) do for each relation r pointed by hindi synset id do if (relation type of r is semantic) do clamp the synset identity linked by relation r in to marathi synset id end if else clamp the synset identity linked by relation r in hindi_synset_id to marathi_synset_id AND manually insert the corresponding lexical element end else end for end if end for

Eg.:Here we have hindi input to English output Marathi word Hindi Word दुकान

दुकान

Take the word **द्वा**.

Consider the id number of the word. Now search for the id number in the English set. If match is found, display the output. Else manually insert the word in the table with an id number.

> ⇔ To find a match between English & Marathi synsets:

for each synset identity marathi_synset_id in Marathi Ontology do if (marathi_synset_id ==english_synset_id) do for each relation r pointed by english synset id do if (relation type of r is semantic) do clamp the synset identity linked by relation r in to marathi_synset_id end if else clamp the synset identity linked by relation r in english_synset_id to marathi_synset_id AND manually insert the corresponding lexical element end else end for

end if end for

Eg: Here we have hindi input to English output Marathi word English Word

दुकान bov _____ Take the word **दुकान**.

Consider the id number of the word. Now search for the id number in the English set. If match is found, display the output. Else manually insert the word in the table with an id number.

3)Uml Diagrams :











Fig.7 SEQUENCE DIAGRAM (for earching operation)

5) Data Diagrams:



6) Interface Design (User Interface):

 The following are the screenshots of the user interface:

 Image: http://localhost:8080/WebApplication1/StoreListing.jsp - Windows Internet Explorer

 Image: http://localhost:8080/WebApplication1/StoreListing.jsp

 File
 Edit
 View

 Favorites
 Image: http://localhost:8080/WebApplication1/StoreListing.jsp

 File
 Edit
 View

 Favorites
 Image: Image:

Choose Category : Mobile 💌

Choose Brand : NOKIA 💌

Available products (with price)are :

Products		Price	
Ν	45	200.0	
E	53	300.0	

Back to Home Page

Fig.9 Front End for SEARCHING

otin Free Provide the the test of test		
E http://localhost.8000/WebApplication1/StoreListing.jsp		
File Edit View Favorites Tools Help		
x Google v 🖓 Search + v 🚳 + 🖶 🖉 Share + 🔊 + 📮 Sidewil		
🔆 Favorites 🛛 🍰 🍘 Suggested Sites 🔹 🙋 Free Hotmail 🙋 Get More Add-ons 🔹		
🔠 🔻 🖉 Project Translation 🦉 http://localhost:8080/We 🗙		
This Project help locate the price of a Product.		
Choose the languange : ○English ⊚ हिन्दी ○ मराथी		
Choose Dept : एलेक्नोनिक 💌		
Choose Category : H_Mobile		
Choose Brand : H_NOKIA 💌		
Available products (with price)are :		
Products Price		
H_N45 200.0		
H E63 300.0		

Back to Home Page

Fig.10 Front end for searching in hindi

Fig.8 ER Diagram

C http://localhost	:8080/Web/	Application1/Product	SearcherByPri	ice.jsp - Win	dows Interi	net Explorer
	p://localhost:8	080/WebApplication1/Prod	uctSearcherByPric	e.jsp		
File Forward w F	avorites Too	ols Help				
x Google		١	/ 🖁 Search 🕶	·	🚱 Share 🔹	🔊 • 🔲 Sidewiki •
😭 Favorites 🛛 😭 (🏉 Suggested	Sites 🔹 🙋 Free Hotmail	🖉 Get More A	dd-ons *		
📲 🔻 🏉 Project Tra	nslation	€ http://localhost	:8080/We X			
This Project help	locate the	products in the spe	cified Price r	ange.		
Choose the langua ©English 0 हिन	nge∶ दी ○मराश	ी				
Enter Price Range	:					
From: 100		To: 300				
Search Cle	ear					
Available products	s (for the sp	ecified price range)are	:			
Price Produ	ct Brand	Category Dept				
200.0000 N45	NOKIA	Mobile Electronic				
300.0000 E63	NOKIA	Mobile Electronic				
					Back	to Home Page
	Fig.11	Front end for l	Price Ran	ge operat	tion	
C http://localhost:80	Fig.11 80/WebAppl	Front end for 1	Price Ran Windows Interr	ge operat 1et Explorer	tion	
/ http://localhost:80	Fig.11 80/WebAppli	Front end for l ication1/translate.jsp - WebApplication1/translate.jsp	Price Ran Windows Interr	ge operat 1et Explorer	tion	
Øhttp://localhost:80 ♥♥♥♥ ₽ http:// File Edit View Favo	Fig. 11 80/WebAppli locahost:8080/v	Front end for 1 ication1/translate.jsp - VebApplication1/translate.jsp Help	Price Ran	ge operat	tion	
<pre> http://localhost:80 file Edt View Favo x Google </pre>	Fig.11 80/WebAppl locahost:8080/i rites Tools I	Front end for l ication1/translate.jsp - WebApplication1/translate.jsp Help	Price Ran Windows Intern earch • 🖉 • 📹	ge operat net Explorer	tion	devili • 🔂 Boolmark
 http://localhost:80 ● http:// File Edit View Favo x Google ★ Favorites ↓ € € € € • € • € • ♦ € • ♦ € • ♦ € • ♦ € • ♦ € • • ♦ € • • • •	Fig. 11 80/WebAppi locahost: 8080/ ittes Tools I	Front end for 1 cation1/translate.jsp - WebApplication1/translate.jsp Help ¥ Free Hotmal Free Hotmal Get	Price Ran Windows Interr earch * 0 🔊 * 📢 More Add-ons *	ge operat et Explorer	tion	dewiki • 🟠 Bookmark
 Attp://localhost:80 ► Edit View Favo X Google ☆ Favorites ☆ Favorites ☆ Project Transle 	Fig. 1 1 80/WebAppfi locahost: 8080/ ites Tools I suggested Sites tion	Front end for l ication1/translate.jsp - WebApplication1/translate.jsp Heb VebApplication1/translate.jsp Heb Heb Heb Heb Heb Heb Heb Heb	Price Ran Windows Interr earch • • Ø • • More Add-ons •	ge operat et Explorer	tion • 💁 • 🗐 5	devili • 🟠 Boolmark
http://localhost:80 File Edit View Favor X Google Favorites Project Transle This Project uses the second secon	Fig. 1 1 80/WebAppli locahost:8080/v locahost:8080/v lites Tools I lituggested Sites tion e Database	Front end for 1 ication1/translate.jsp - VebApplication1,translate.jsp telp VebApplication1,translate.jsp telp VebApplication1,translate.jsp vebApplication2,translate.jsp vebApplication2	Price Ran Windows Interr earch • • 🗭 • 🖨 More Add-ons • e x	ge operat et Explorer • ② Share Translation.	tion • هِ • الع	dewiki • 🏠 Bookmark
http://localhost:80 マ を http:// File Edit View Favo X Google 文 Google マ 使 Project Transla This Project uses th Choose the source la ③ English ○ 『たテ라	Fig. 1 1 80/WebAppli locahost: 8080/v locahost: 8080/v lites Tools I suggested Sites tion e Database nguange : 0 मराथी	Front end for 1 ication1/translate.jsp - VebApplication1,translate.jsp telp	Price Ran Windows Interr earch • • • • • • • More Add-ons • • • x	ge operat et Explorer • ② Share Translation.	tion • هِ • الع	devili • 🏠 Boolmark
http://localhost:80 Image: Straight of the straig	Fig. 1 1 80/WebAppli locahost: 8080/v intes Tools I suggested Sites tion ee Database nguange : ्रमराथी	Front end for 1 cation1/translate.jsp - telp telp Free Hotmal @ Get Mttp://locahost:8080/W driven approach for S List Synonyms	Price Ran Windows Interr earch • 🛷 • d More Add-ons • e X	ge operat et Explorer	· D s	dewiki • 🟠 Bootmark
http://localhost:80 v v intervention file Edit View Favo x Google favorites favorites favorites intervention file Project Transle Offerfile Choose the source lat o English Offerfile Enter the text: garmene Cloth Clo	Fig. 1 1 80/WebAppl locahost: 8080/4 locahost: 8080/4 lites Tools I liuggested Sites tion ne Database nguange : ा मरायी	Front end for I ication1/translate.jsp - WebApplication1/translate.jsp Heb Free Hotmal @ Get Chttp://localhost:8000/W driven approach for S List Synonyms	Price Ran Windows Interr earch • • • • • • • More Add-ons • • • • • • • • • • • •	ge operat	v Dr. a	devili • 🟠 Bookmark
http://localhost:80 Image: Straight of the straig	Fig. 1 1 80/WebAppli locahost: 8080/v ittes Tools I iuggested Sites tion nguange : ा मरायी	Front end for 1 cation1/translate.jsp - WebApplication1/translate.jsp Help Yeb Free Hotmal @ Get Help Help List Synonyms	Price Ran Windows Interr each • () • (More Add-ons • e X	ge operat	· Dr . s	dewiki • 🛣 Bookmark
Ittp://localhost:80Itte:/localhost:80 </th <td>Fig. 1 1 80/WebAppi locahost: 8080/ ites Tools I iuggested Sites tion e Database nguange : 0 मराथी nt t t</td> <td>Front end for 1 cation1/translate.jsp - Web4pplcation1/translate.jsp Web4plcation1/translate.jsp Web4plcation1/translate</td> <td>Price Ran Windows Interr earch * 0 @ * 4 More Add-ons * e X</td> <td>ge operat</td> <td>· Dr . s</td> <td>dewiki • 🕻 Bookmark</td>	Fig. 1 1 80/WebAppi locahost: 8080/ ites Tools I iuggested Sites tion e Database nguange : 0 मराथी nt t t	Front end for 1 cation1/translate.jsp - Web4pplcation1/translate.jsp Web4plcation1/translate.jsp Web4plcation1/translate	Price Ran Windows Interr earch * 0 @ * 4 More Add-ons * e X	ge operat	· Dr . s	dewiki • 🕻 Bookmark
http://localhost:80 File Edit View Favorites Favorites Favorites Favorites Favorites Favorites Favorites Favorites Favorites File Favorites Favorites Favorites Favorites Favorites Favorites Favorites Enter the rescurred language Farglish Farglish Farglish Cloth Clothing List of Sym Clothing List of Sym Farglish Farglish Farglish Translate Entered Te	Fig. 1 1 80/WebAppli locahost: 8000/ intes Tools I suggested Sites tion nguange : ा म ा म ा स स म साथी सराथी	Front end for 1 cation1/translate.jsp - WebApplication1/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/translate.jsp WebApplication2/tran	Price Ran Windows Interr earch • • • • • • • More Add-ons • • • • • • • • • • • • • • • • • • •	ge operat	· Dr . s	devili • 🟠 Boolmark

Fig.12 Front End for Inter-language Translation

V. PROJECT TESTING

Sr. No.	1
Test Name	VIEW TRANSACTIONS: The user should be able to see all the previous transactions he has performed from the current account. He can also select a particular transaction from the list to view the complete details.
Test Objective	Client Testing(functional)
Test Configuration	 Installation of application Initiation of application Display of the first screen
Procedure	Launch the application.
Action	Response
 Launch the application Select VIEW TRANSACTI ON option from the screen 	 The transaction list is displayed Upon selection of a particular transaction the full detail should be displayed
Expected Results	If there are any transactions performed then a list of all will be displayed else an error message stating "No Transactions" will be displayed.
Actual Result	AS EXPECTED
PASS/FAIL	PASS
Sr. No.	2
Test Name	CANCEL TRANSACTION: This screen allows a user to place a request for the cancellation of a particular transaction and only if the customer selects to cancel, a

	request the transaction is cancelled.
Test Objective	Client Testing(functional)
Test Configuration	 Installation of application Initiation of application Display of the first screen
Procedure	Launch the application.
Action	Response
 Launch the application Select CANCEL TRANSACTI ON option from the screen 	 Confirmation displayed Cancel transaction if both the participating entities have placed a cancellation request
Expected Results	The cancellation-accepted request message will be displayed
Actual Result	AS EXPECTED
PASS/FAIL	PASS

VI. REFERENCES

Vossen, Piek (ed.) 1999. EuroWordNet: A Multilingual Database with Lexical Semantic Networks for European languages. Kluwer Academic Publishers, Dordrecht.

Christodoulakis, Dimitris N. 2002 . BalkaNet: A Multilingual Semantic Network for Balkan Languages. EUROPRIX Summer School, Salzburg Austria, September 2002.

Ramanand, J., Akshay Ure, Brahm Kiran Singh and Pushpak Bhattacharyya. Mapping and Structural Analysis of Multilingual Wordnets. IEEE Data Engineering Bulletin, 30(1), March 2007.

Deeshmukh, Amit A., Mansi Mohan, A. Amrita, and K. P. Ray. "Broadband Proximity Fed Equilateral Traingular Microstrip Antenna." In *2012 International Conference on Advances in Computing and Communications*, pp. 263-266. IEEE, 2012.

Sinha, Manish, Mahesh Reddy and Pushpak Bhattacharyya. 2006. An Approach towards Construction and Application of Multilingual Indo-WordNet. 3rd Global Wordnet Conference (GWC 06), Jeju Island, Korea, January, 2006.

T. Berners-Lee, J, Handler, O Lassila. "The Semantic Web," Scientific American, May 2001.

Iyengar, Makrant M., and Mansi M. Iyengar. "SEMANTIC MATCHING USING ONTOLOGY IN MULTILINGUAL ENVIRONMENT (WITH GUI & IMPLEMENTATION)." *International Journal of Global Technology Initiatives* 2, no. 1 (2013): B20-B27.

M. Uschold and R. Jasper, "A Framework for Understanding and Classifying Ontology Applications," Proceedings of IJCAI Workshop on Ontologies and Problem-Solving Methods, Aug. 1999.

Agirre E. and Rigau G. (1996) Word Sense Disambiguation using Conceptual Density. COLING, Denmark

Ide. Nancy (1999) Parallel Translation as Sense Discriminator. In Proceedings of SIGLEX, University of Maryland, College Park, USA

Deshmukh, Amit A., Mansi Mohan, Raj Shah, and Prateeksha Runwal. "Analysis of Broadband Proximity Fed Gap-coupled C-shaped Microstrip Antennas." (2014).

Niles, I., & Pease, A. 2001. Toward a Standard Upper Ontology, in Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), Chris Welty and Barry Smith (eds.).

Niles, I., and Pease, A. 2003. Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology, Proceedings of the IEEE International Conference on Information and Knowledge Engineering, pp 412-416.